THE RELATIONSHIP OF SELF-PACED INDIVIDUALIZED INSTRUCTION TO PUPIL ACHIEVEMENT WHEN MEASURED BY POOLING THE PROBABILITIES OF SEVERAL INDEPENDENT SAMPLES

By

PAUL IVAN JOHNSON

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF
THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1979

ACKNOWLEDGMENTS

Grateful acknowledgment is made to the many individuals who made this study possible. Sincere appreciation is extended to Dr. Ralph B. Kimbrough (chairman), Dr. Robert S. Soar, Dr. Michael Y. Nunnery, Dr. Herbert Franklin, and Dr. Paul S. George who made many valuable suggestions as members of the dissertation committee.

To Dr. Ralph B. Kimbrough, Dr. Robert S. Soar, and Dr. Michael Y. Nunnery the writer owes a special debt of gratitude for their advice which was always clear and directly to the point.

Finally, the writer would like to acknowledge his indebtedness to his wife, Dona, without whose understanding, encouragement, and inspiration, this dissertation probably would not have been completed.

TABLE OF CONTENTS

TABLE OF CONTENTS (con't)

Abstract of Dissertation Presented to the Graduate Council of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

THE RELATIONSHIP OF SELF-PACED INDIVIDUALIZED INSTRUCTION TO PUPIL ACHIEVEMENT WHEN MEASURED BY POOLING THE PROBABILITIES OF SEVERAL INDEPENDENT SAMPLES

By

Paul Ivan Johnson

June 1979

Chariman: Dr. Ralph B. Kimbrough
Major Department: Educational Administration & Supervision

The problem of inconsistent results in educational research has been pervasive in the education profession. Administrators and other educational decision makers have, under the pressure of economic, political, and social forces, often been pressured to adopt educational innovations without prior validation through empirical research. Self-paced individualized instruction was an example of an instructional innovation which has been widely adopted by school decision makers without such validation. Research results reported on self-paced individualized instruction have presented an inconsistent and inconclusive data base from which to formulate an objective decision.

Some researchers have reported that self-paced individualized instruction is superior to traditional classroom instructional practices. Other researchers have reported

the opposite result and yet others have reported "no difference" in the two methodologies when those methodologies were comapred by their effect on student achievement. The concept of secondary analysis promised a solution to the problem of inconsistent results thereby making it possible for decision makers to base their decisions on reliable research. In secondary analysis, the researcher gathers data from previously completed research studies. The researcher reorganizes and reanalyzes the data through a variety of possible techniques.

In this study, the aggregate chi-square statistical procedure was applied to the reported results of a preselected set of independent studies. By pooling the statistical results of 24 independent samples taken from 11 independent research studies, it was shown that traditional classroom practices are superior to self-paced individualized instruction when pupil achievement is used as the criterion. The derived statistic was aggregate χ^2 (48)= 101.9880, p <.002.

The above derived statistic shows that self-paced individualized instructional practices are not superior to traditional instructional practices. In fact, the reverse could be implied; that traditional instructional practices are superior to self-paced individualized instructional practices. The probability of .002 for

this combined significance test indicated that in 998 of every 1,000 cases this result would be replicated. In other words, it was very unlikely that this result came by chance.

CHAPTER I

INTRODUCTION

Need For The Study

The school administrator has, among others, the functions of decision maker and instructional leader. The accountability movement has brought considerable pressure on the local school administrator to justify the programs and practices which have been and are being used in the classrooms. The pressure exerted by the accountability movement has caused administrators to scrutinize more carefully the instructional practices used in their schools. In the past, many instructional innovations have been advanced and adopted because of emotional or philosophical commitments made by educational leaders and not because of evidence garnered through the empirical research process.

The results of educational research have also provided a tenuous base upon which to build classroom practices. One of the reasons for this was that classroom practices have often been predicated on the basis of one or two research studies or on no research at all. Very often one researcher's conclusions have contradicted another's with all of this producing a state of confusion among educational decision makers and practitioners.

Pressure from the demands of the accountability movement, coupled with the historically capricious results of educational research and with the tendency of educators to emotionally commit themselves to new ideas, has resulted in many educational methodologies having been adopted apart from any consistent empirical validation. Self-paced individualized instruction was one such methodology. Some researchers have reported that self-paced individualized instruction was superior to traditional classroom practices. Others have reported the opposite result and many more have reported that there was no difference in the two methodologies in terms of enhancing student achievement.

Obviously, new or better techniques were needed to help resolve the issue. Several researchers (Gage, Note 1; Glass, 1976; Light & Smith, 1971) have suggested secondary analysis techniques as a means of resolving the dilemmas that prevailed in educational research results. Secondary analysis was needed in the case of self-paced individulaized instruction because of two inherent problems that plagued self-paced individualized instruction research studies.

The problems of low expected relationships and small sample size have been pervasive problems in all classroom research as well as in the case of self-paced individualized instruction. The school classroom has been noted for its multiplicity of variables. All these variables could have an effect on the results of instruction. Therefore, the expected relationship of any one variable (whether textbook,

teacher style, room temperature, instructional method, etc.) to a learning outcome (grade or test score) was automatically a low relationship. Generally, no one variable would have a high correlation with any other variable. A second problem plaguing all classroom research has been that of small sample size. It has been difficult to generalize the results of classroom research to larger populations because the original studies were carried out with small samples. There were other problems generated by the use of small samples and these were discussed later in this study.

Studies completed on self-paced individualized instruction seemed to fit the pattern which had been evident in other educational research. Secondary analysis techniques were needed in order to resolve the inconsistent pattern of research results coming from studies done on self-paced individualized instruction. A secondary analysis technique called the aggregate chi-square was suggested by Gage (Note 1) as a possible solution to the pervasive problems affecting classroom research. The following sections are a more detailed discussion of the accountability movement and its effect on classroom practices, the weakness of past educational research, the tendency of educators to commit themselves to innovative ideas before or without empirical validation, the need for secondary analysis in educational research, the two pervasive problems in educational research, the potential solution, and why administrators need a more sound base from which to make decisions.

Accountability Demands Justification of Classroom Instructional Strategies

The "formidable force" (Sciara & Jantz, 1972, p. 3) of accountability has been a part of the educational scene since 1963. Darland (1970) called it a "national crisis" and asserted that "the American teacher has become a most likely candidate for scapegoat of the 1970's (p. 41).

But, according to Briner (1969), it is the school administrator who has been and will continue to be the scapegoat.

> Teachers, parents, and others apparently don't want to be fully accountable for improving education. They tend to be consumed by their typical social and professional roles. On the other hand, administrators must be directly accountable. It is their essential reason for being. They are the ones upon whom the hands of approval or disapproval will be laid. Accordingly, their roles require orientation and commitment to educational success and the elimination of failure. In their performance, they can be expected to explain both success and failure; they must be capable of proposing educational improvements to the satisfaction of students, teachers, parents, and others. (p. 205)

The move to accountability in education has produced an "imperative" (Hostrop, Mecklenbarger, & Wilson, 1973, preface) force upon educators (whether teacher or administrator) to show "proof of results" (Lessinger, 1971, p. 13). In addition, Lessinger asserted that general results will not be proof enough. "We must go beyond such general outlines of general results and find out what specific factors produce specific education results" (p. 13).

Authors of articles and books on educational accountability have stressed two important factors. First, that educators must be able to show which instructional methods and programs are successful and second, (and more specifically) success of educational programs will be measured in terms of pupil achievement.

> Thus, for our purposes, classroom activities will be deemed successful if they induce desired changes in pupils. Such a position requires us--eventually--to validate the practices of teaching with product variables [product variables are evidences of change in students as a result of their involvement in the classroom]. (Dunkin & Biddle, 1974, p. 48)

The forces of accountability have demanded and will continue to demand specific results from specific programs (Berry, 1977, p. 4). Educators have not been and are still unable to document with empirical evidence the merit of classroom practices. The "bottom line" in education is proof of results. Educators have been unable to offer the proof demanded.

The Weakness of Educational Research

In 1953, the Committee on Teacher Effectiveness of the American Educational Research Association reported the following conclusions.

> The simple fact of the matter is that, after 40 years of research on teacher effectiveness during which a vast number of studies have been carried out, one can point to few outcomes that a superintendent of schools can safely employ in hiring a teacher or granting his tenure, that an agency can employ

> in certifying teachers, or that a teacher-education faculty can employ in planning or improving teacher-education programs. (p. 657)

Twenty-one years later, Dunkin and Biddle (1974) concluded that "those who are seeking simple answers to the problem of teacher effectiveness are only slightly better off today than they were twenty years ago" (p. 16). Two years after Dunkin and Biddle, the following observation was made.

> The status of numerous innovative instructional strategies is extremely tenuous, with many being perpetuated by a false sense of empirical security. Teachers are implementing a variety of techniques which they assume to be superior but which have not been proven in practice, at this point in time. (George & Maxwell, 1976, p. 56)

Education has had a history of weak empirical validation of its programs and practices. "Educational research typically follows innovation and hence has little constructive effect on educational practice" (Good, Biddle, & Brophy, 1975, p. vii).

The preceding comments were made by professional educators all of whom are deeply committed to achieving quality in education through empirical research. If, in their judgment, educational practices have been and still are based on a tenuous empirical foundation, is it any wonder that the educational establishment has been forced into a defensive position by the accountability movement.

The "Commitment" of Educators to Individualized Instruction

Individualized instruction is an educational practice that needs empirical justification. Educators have made a heavy commitment to individualized instruction. Dunkin and Biddle (1974) defined commitment as "advocating innovative ideas for improving education" when those ideas are "attractively argued but unsupported by data" (p. 51).

Individualized instruction became one of the basic commitments of the education profession under the pressure of accountability. Morris (1971) in commenting on "ten things accountability will require" stated that one of those requirements will be a commitment to individualized instruction.

> Education must become, not only in theory but in fact, child centered. We will be forced to write programs for each child based on extensive results of highly sophisticated diagnostic instruments. (p. 326)

Another writer (Davies, 1970) supported Morris' prediction.

> Teachers and all school personnel are involved in the search for answers to critican questions such as these:
>
> How do we move from a mass approach to teaching and learning to a highly individualized approach?
>
> How do we go about the "simple" task of treating each child as an individual human being? (p. 129)

Davies' and Morris' comments were made in response to the accountability movement which was beginning to have a strong impact (Hostrop et al., preface) on education at the time of their writings.

Professional educators have more recently affirmed their commitment to individualized instruction. Saylor (1977) wrote in the guest editorial of the January issue of Educational Leadership the following statement:

> Judgments about the quality of education must be rendered in terms of how well the school is developing the respective set of talents, capabilities, and potentialities of each student for living a life of personal satisfaction and compassionateness in our society. (p. 245)

The subsequent February issue of Educational Leadership was devoted to the topic of individualized instruction. Anderson, in the guest editorial to that issue, wrote that "the resolve of most educational leaders is to pursue the path toward individualization" (p. 324). It was apparent that educators have made the commitment to individualized instruction. Have the results of classroom research supported this commitment? It was unlikely that a commitment to individualized instruction was justifiable.

Educational Leadership, as noted above, devoted its February, 1977, issue to individualized instruction. Following are some of the comments by different authors in that issue.

> Individualized instruction is an excellent example of the truth that educational innovations should be tried cautiously, with proof demanded that they actually produce better results. (Weber, p. 328)

> I am pessimistic that the concept will survive [that is, the concept of mastery of learning] unless the practice of individualized instruction proves to be more effective. (Block, p. 341)

> The results have been conclusive. Implementation of the ACIL [Arizona Consortium for Individualized Learning] process works for students, teachers, administrators, and parents. (Webb & Howard, p. 356)

> In a time which has been described as education's Age of Individualized Instruction there is a pressing need for a comprehensive evaluation of all instructional strategies [instructional strategies pertaining to carrying out the individualized instruction concept]. (George & Maxwell, p. 57)

Secondary Analysis: A New Trend in Educational Research

The pressure of the accountability movement and the historical unreliability of educational research has caused educators to seek new ways to conduct research on instructional strategies.

> In educational research, we need more scholarly effort concentrated on the problem of finding the knowledge that lies untapped in completed research studies. . . . The best minds are needed to integrate the staggering number of individual studies. This endeavor deserves higher priority now than adding a new experiment or survey to the pile. (Glass, 1976, p. 4)

Glass (1976) was referring to a relatively new approach (of which he was a forerunner) in educational research. The new approach has been called "secondary" or "meta" analysis (Glass, 1976, p. 3). The basic thrust of secondary analysis is that the researcher analyzes previously done analysis on a given topic or variable (analysis of analysis).

Two conditions in educational research have led to the need for secondary analysis. One condition is that there is a proliferation of studies on a wide variety of educational topics, but little has been done to integrate the results of the many individual studies.

A second condition is that the "fragile and confusing findings" which has been characteristic of all educational research (Glass, 1976, p. 3). "Where ten studies might suffice to resolve a matter in biology, ten studies on computer assisted instruction or reading may fail to show the same pattern of results twice" (p. 3).

The key problem was that of integrating the findings. How shall educational researchers integrate the varied results? There have been at least three approaches to the problem of integration of research findings.

One attempt at integration has been made by simply discarding all of the studies of a given set which have not been done with acceptable designs or systems of statistical analysis. This method of integration leaves the researcher with a set of studies which are technically near perfection in terms of design and analysis. Glass decried this method and commented "that this approach takes design and analysis too seriously" (Glass, 1976, p. 4). Glass further commented that "eliminating the 'poorly done' studies is to discard a vast amount of data" (p. 4).

A second method of integration has been to collect a significant number of studies done on a given variable and compare the reported statistics in terms of the direction of significance. Dunkin and Biddle (1974) used this method in their comprehensive review of the research on teaching. The inconclusive results reflected in the Dunkin and Biddle book are what led Gage (Note 1) to pursue a secondary analysis of the same data.

Secondary analysis, then, is the third method which has been used to integrate research data. Glass (1976); Light and Smith (1971); and Gage (Note 1) have suggested various approaches and statistical techniques for secondary analysis. The method of secondary analysis used in this study follows the aggregate chi-square described by Gage (Note 1). The use of the aggregate chi-square procedure was especially appropriate because of two problems that are inherent in educational research.

Two Pervasive Problems of Classroom Research

Why have educators been unable to substantiate classroom practices through empirical research? Gage (Note 1) identified two factors that continually plague educational research. One factor was that of the "expected relationship" of any single teaching variable to the effect on students who are exposed to that variable. A second factor was the effect of small sample size (Gage, Note 1, pp. 8-9).

Gage has pointed out that, first, the relationship of "any single variable of teacher behavior" to that of pupil achievement was "probably low"

> On the face of it, the teaching-learning process is so complex that any single variable of teacher behavior should have only a low correlation (ranging from about $\pm$ .1 to about $\pm$ .4) with student achievement or attitude. (Gage, Note 1, p. 8)

A second problem that plagued educational researchers was the problem of small sample size. Most educational research was based on relatively small numbers of teachers.

> For a sample of the median size, namely, 15 teachers, it is necessary that a correlation coefficient equal .51 if it is to be significant at the .05 level. The coefficient must equal .64 if it is to be significant at the .01 level. (Gage, Note 1, p. 8)

When the two problems, low expected relationships and small sample size, were combined, it was not unexpected that most of the studies in educational research would not attain statistical significance.

Is it possible to gain significant knowledge from classroom research that has the inherent problems of low expected relationships and small sample size? Can classroom research, faced with the two inherent problems just described, justify the programs and practices being used in American classrooms through empirical research (specifically, self-paced individualized instruction)?

A Potential Solution

Gage followed his description of the problems with a potential solution. It is a method for testing the significance of combined results. Gage applied the aggregate chi-square model to five clusters of studies that had been done on different teacher behavior variables (Gage, Note 1, p. 13). By applying the aggregate chi-square model, Gage was able to show that teacher variables (such as "praise") had significant relationships to student achievement. The importance of those findings is that in most previous research on those same variables, those variables did not prove to have a significant relationship to student achievement. Gage summarized the importance of his findings in the following paragraph:

> These tests of combined results do indeed reveal some significant relationships between types of teacher behavior and student achievement or attitude. Thus, the results suggest that seeking such process-product relationships is not altogether fruitless. They bear out the widely held and hard to surrender intuition that how teachers behave makes a difference in what students learn. (p. 15)

The method used by Gage and the resultant findings indicate that research on teaching may not be as ambiguous as previously believed.

Because the results of independent research studies done on self-paced individualized instruction presented the same ambiguous picture as studies done on many other instructional variables, it was appropriate to apply the

aggregate chi-square procedure in a secondary analysis of self-paced individualized instruction.

Administrator Need for Objective Data Regarding Self-Paced Individualized Instruction

Individualized instruction (specifically, self-paced) was a classroom instructional practice widely adopted and touted as the answer to the demands of accountability. Individualized instruction was also advanced as the means through which the individual could reach his maximum potential. Furthermore, it was written in a leading educational journal that "the resolve of most educational leaders is to pursue the path toward individualization" (Anderson, 1977, p. 324).

Did the accumulated empirical evidence from research done on self-paced individualized instruction justify such a strong commitment? The evidence did not justify such a commitment. However, the empirical evidence did not preclude such a commitment either. As was mentioned earlier in this study, the research on self-paced individualized instruction was inconsistent and therefore inconclusive.

In the face of the pressures from educational consumers (accountability) and from educational professionals, an administrator needed a more objective data base from which to generate a viable decision. Decisions involved with an innovation such as self-paced individualized instruction were not insignificant in terms of finances, administrator time, teacher time, in-service, curriculum changes, hardware,

software, etc. Stronger empirical evidence regarding self-paced individualized instruction was needed.

Statement of the Problem

The problem of this study was to determine the effectiveness of self-paced individualized instruction by applying the aggregate chi-square statistical procedure to the results of several, previously done, independent studies which have analyzed self-paced individualized instruction in its relationship to student achievement. What would be the result when these several independent results were pooled to derive a single statistic?

Delimitations and Limitations

In a study of this nature there are certain constraints on any results that might be obtained. The following delimitations and limitations are explanations of the constraints applicable to this study.

Delimitations

1. The purpose of this study was to investigate the aggregate main effect of self-paced individualized instruction on pupil achievement. Consequently, interaction effects from within or across the individual studies were not investigated.
2. The test of combined significance bore upon a question being asked about the cluster of studies as a whole and consequently did not bear

on any of the individual studies included in the cluster. That is, the combined result does not change and cannot be construed to change the previous results of the independent studies.

3. It was beyond the scope of this study to make an assessment of the experimental studies analyzed except that they met the criteria outlined in the procedures section.
4. The set of 11 studies used to obtain the pooled statistic was not exhaustive (it did not include all possible studies done on that subject or that grade level).

Limitations

1. The results of this study are not decisive or final in the case of self-paced individualized instruction and its effect on pupil achievement. Replications of this type of study, other secondary analysis techniques, or other original research would be necessary to support the conclusions reached in this study.
2. The results of this study were not generalizable beyond self-paced individualized instruction at the high school level in mathematics oriented subjects.

Definition of Terms

For the purposes of this study, the following terms were used as defined below.

Control group. In experimental studies, "the control group does not receive the experimental treatment [independent variable]" (Huck, Cormier, & Bounds, 1974, p. 245). The control group must be equivalent to the experimental group in respect to all crucial variables. Equivalency of control and experimental groups is usually achieved by random assignment of subjects to groups. Other equalization techniques are often necessary to support random assignment and to insure equalization of control and experimental groups.

Dependent variable. "The researcher in experimental research must first identify those dependent variables which will, taken together, make a reasonable test of the independent variable? (Fox, 1969, p. 460). The selection of dependent variables is determined by the research problem, stated hypothesis, and the class or type of dependent variable or variables being measured. Some dependent variables may be measures of achievement (as one class or type) while others may be associated with attitude. Some dependent variables measure long-term results while others measure short-term results. The dependent variable must be consistent with the purposes of the research problem, measure results that correspond to the hypothesis being tested, and be measured with instruments that are valid for the particular subjects being

used in the study. The goal of the researcher is to show a casual relationship between the independent and dependent variable or variables.

Experimental group. In experimental studies, "the group that receives the treatment [independent variable] is called the experimental or treatment group" (Huck et al., 1974, p. 245).

Independent variable. The independent variable is the condition or conditions that an experimenter can manipulate. The experimenter attempts to show a casual relationship between the independent variable and some outcome measure (dependent variable). The independent variable is sometimes called the "experimental, treatment, or intervention variable" (Huck et al., 1974, p. 224).

Individualized instruction. Most authors agree that there is no one definition for individualized instruction. "The only common universal attached to the term is that students generally will be able to proceed at their own pace in some areas" (Good et al., 1975, p. 169).

Process variable. Process variables are the conditions in a given classroom that would be expected to effect the performance of the pupils in that classroom. A teacher's personality, the textbooks, audio-visual aids, and grouping practices are examples of process variables.

Process-product research. Process-product research is the attempt of researchers to establish casual relationships between specific process variables and specific product variables.

Product variable. Product variables are the outcome measures associated with given classroom groups or with individual pupils. A pupil's test score or a group mean derived from a set of test scores are examples of product variables. Test scores used as outcome measures are usually a measurement of pupil achievement or pupil attitude toward school.

One and two-tailed tests. The definition offered by Huck et al., (1974, pp. 45-46) was accepted for this study and it is as follows:

> When using some tests of significance, the researcher must also decide whether the test will be a one-tailed test or a two-tailed test. A two-tailed test is sensitive to significant differences in either direction (i.e., greater and less); the one-tailed test is sensitive to differences in only one direction (i.e., greater or less). Suppose, for example, that a researcher compares achievement test scores of a group of students exposed to a new method of instruction to the scores of another group instructed by the traditional method. If a two-tailed test is used to compare the scores of both groups, the researcher can answer two questions: (1) Do students under the new method score significantly higher? (2) Do students instructed with the traditional method score higher? On the other hand, if a one-tailed test is used, the researcher can answer only one question: (1) Do students under the new method score significantly higher?
>
> Also, if differences are found to be significant at a certain level of significance with a one-tailed test, the same difference with a two-tailed test would be significant at a level of significance twice as large as that used with a one-tailed test. For example,

> if the researcher found a significant difference at the .025 level with a one-tailed test, the same data used with a two-tailed test would be significant only at the .05 level.

Traditional instruction. The researchers whose data were pooled in this investigation used the term "traditional" to define a mode of instruction which contrasted to individualized instruction. "Traditional" was synonymous with large group instruction, the lecture-discussion mode, and teacher-paced learning environments.

Procedures

The general approach was to use the procedures developed by Gage (Note 1) to reanalyze the relationship of individualized instruction to pupil achievement. The studies identified by Schoen (1976) were selected for analysis. The aggregate chi-square statistic was applied to the data as demonstrated by Gage.

Source of Data

The procedure and rationale for selecting the source of data for this study followed closely that of Gage (Note 1). Gage chose a process variable which had been reported on in the Dunkin and Biddle text (1974) and which had inconclusive results when a number of independent studies, reporting on the same process variable, were compared. For example, teacher praise was one process variable on which Dunkin and Biddle gathered evidence. The studies reported on by Dunkin and

Biddle gave inconsistent results and presented a confusing picture as to the effect of teacher praise on learning outcomes.

Gage took the same set of studies collected by Dunkin and Biddle and applied the aggregate chi-square statistic to pool the independent probabilities into an aggregate statement of significance. As was reported earlier, the pooled statistic revealed a consistent relationship between the use of praise and pupil achievement.

The purpose of this study was to ascertain the relationship of individualized instruction to pupil gain in academic achievement by using the pooled probabilities of several independent samples. A survey of the literature revealed that several reviews of research had been published with regard to individualized instruction. Schoen's (1976) review was the one chosen. In Schoen's review, one researcher's results favored the experimental group (individualized instruction), three other researchers reported results favoring the control group (traditional instruction), and eight other researchers reported results as not significant (Schoen, 1976, p. 354).

Requirements and Statistical Procedures

This section is an explanation of the requirements and procedures (Fisher, 1948, pp. 99-101) which are necessary when applying the aggregate chi-square model to a set of data. As written previously, the aggregate chi-square model was

employed to reanalyze data from several selected studies. By following the steps outlined below, the reader should be able to duplicate this type of study.

Independence of the separate results being combined. This means that each result used was obtained from a different sample and that only one set of results for any single sample of individuals was usable.

Some experimental studies were done using several schools and/or teachers. When one or more schools or teachers were treated as separate experimental units in a given study, then each school and/or teacher was a source of an independent result from an independent sample. For example, if in a certain study, teacher A, teacher B, and teacher C were each assigned to a control group and to an experimental group, the statistical results were independent for each teacher. The sample of students for teacher A was separate from each other set of control and experimental groups. Likewise, the sample associated with teacher B and the sample associated with teacher C was separate and independent of each other.

For another example (which is a somewhat different case than the one noted above), if in a given study there is more than one dependent variable (criterion measure), the independence of the statistical results can be maintained by reporting the least significant chi-square, the most significant chi-square, and a mean (mean of the two or more results chi-square (Gage, Note 1, pp. 14 & 33).

Rational basis for combining the studies. All of the studies must investigate the same variable. An example would be a set of studies all of which investigated the relationship of time in class to student achievement. In this study, the rational basis was the use of the individualized instruction approach and its effect on student achievement. The rational basis for this particular study was further defended in Chapter II.

All results in the set being analyzed must be included. An objective means of achieving the above criterios was to set specific limits on what studies were included. The researcher could use criteria such as subject matter, grade level, studies done only within a given span of time (1960-1970, for example), and/or a criterion such as experimental studies as opposed to field studies.

Another procedure (the procedure chosen for this study) was to choose all of the studies included in some review of research as when Gage (Note 1) chose only the studies reported by Dunkin and Biddle (1974) in their review of classroom research. The important consideration was that the researcher avoid the bias that would occur if he chose only those studies which offered near significance.

Knowledge of the exact one-tailed probability value associated with each result to be combined. Researchers, in many cases, have reported statistical measures such as *F* ratios, *t*-ratios, and chi-squares. In addition, those statistics have been reported as being either significant (at a

given level of probability) in favor of the control group, significant in favor of the experimental group, or not significant (NS).

Some studies were set up on the basis of an hypothesis which required a two-tailed test for significance and some hypotheses required a one-tailed test for significance.

If the hypothesis was non-directional, the exact probability value associated with the reported statistic was divided by two ($p/2$). For example, if the hypothesis required a two-tailed test and the reported F for that hypothesis was associated with a .30 probability value, then the exact one-tailed probability value is .15 (.30/2). When the reported probability value was associated with a one-tailed test, dividing by two was unwarranted.

Determination of the direction of the statistical results. Did the reported statistic favor the control group or the experimental group? The researcher indicated which group was favored when the statistic was significant. In many cases, the reported statistic was reported as not significant. When this happened, the actual results (the group means) were obtained from the body of the original research report. The group means indicated which group did better on the criterion measure or dependent variable.

If the reported statistical result and the exact one-tailed probability associated with that result were in the same direction as the majority of cases in the set of studies whose statistical results were being pooled, then the exact

one-tailed probability value was taken directly to the chi-square table for two degrees of freedom developed by Gordon, Loveland, and Cureton (1952).

The exact one-tailed probability was subtracted from one (1-p) before going to the chi-square table with two degrees of freedom if it was not in the same direction as the majority of cases in the set of studies whose statistical results were being pooled.

Adjustment of aggregate one-tailed probability to a two-tailed probability. The probability derived from the aggregate chi-square statistic had to be doubled. The use of one-tailed probabilities to derive the aggregate statistic and probability was inconsistent with the assumption that the procedure allows for a result in either direction. This inconsistency is easily remedied by doubling the probability derived from the aggregate chi-square statistic (Gordon et al., 1952, p. 315).

Summary of Procedures for Treatment of the Data. The following summary was presented to help clarify the individual steps in the aggregate chi-square procedure.

1. The exact one-tailed probability associated with the reported statistic was found for each independent sample.
2. The direction of the result was determined. Was it in the same direction as the majority of the cases included in the set of studies being pooled?

If it was in the same direction as the majority, the exact one-tailed probability was taken directly to the chi-square table with two degrees of freedom (Gordon et al., 1952). If the result was not in the same direction as the majority, the probability value was subtracted from one (1-p). When the result was generated from a two-tailed test, the probability was divided by two (p/2) before going to the extended table of chi-square developed by Gordon et al., 1952.

3. All of the independent chi-square statistics with two degrees of freedom were added together to obtain the aggregate chi-square statistic.
4. The aggregate chi-square statistic was taken to a regular chi-square table of values and the appropriate probability value was read from that table.
5. The derived probability was doubled to correspond to the two-tailed assumption which allowed for a result in either direction.

Organization Of The Study By Chapters

Chapter I was an introduction to the study. Topics discussed in Chapter I were need for the study, the problem, delimitations and limitations, definition of terms, and procedures.

Chapter II is a discussion of the rationale for choosing the set of 11 research reports included in this study which were all from the Schoen (1976) review.

Chapter III is a presentation of each of the 11 individual research reports (in abstract form) with their statistical results and the aggregate statistic derived by pooling the individual statistics.

Chapter IV is a summary of this research study as well as a discussion of the conclusions, implications, and suggestions for further research.

CHAPTER II

SELECTING A SET OF STUDIES

Introduction

To accomplish the purpose of this study, a set of independent samples had to be identified. Reading and mathematics have traditionally been the most individualized of all school subjects. Therefore, a set of studies was sought where each investigator had done an experiment on individualizing and where all of the experiments in the set were on reading or math, but not both.

An additional requirement was that self-pacing had to be an identifiable feature of each study. Self-pacing has been identified as the only common denominator among various individualized programs (Good et al., 1975, p. 169 and Miller, 1976, p. 345).

Miller's Review

Miller (1976) produced a review of research on individualized mathematics programs. One hundred and forty-five studies were included in Miller's review. The studies included were done over a wide range of grade levels (kindergarten through college level) and over a wide range of time (17 of the studies were done before 1960).

Primarily, because of these wide ranges in grade levels and time, Miller's review was not used. An additional reason was that the sources of data were so varied that a great amount of time and expense would have been necessary to sort out and track down a representative set of studies.

Hirsch's Review

Hirsch (1976) reviewed research on individualized instruction in secondary mathematics. Thirty-three studies were reviewed. The grade level range of the studies was from grade seven through grade twelve. The time range was from 1967 through 1974.

Although Hirsch had narrowed the field considerably from that of Miller (1976), the grade level span was not ideal. Time and expense were also a factor in deciding against using Hirsch as a source.

Schoen's Review

The set of studies chosen was taken from Schoen's (1976) review which he titled "Self-paced Mathematics Instruction: How EFfective Has It Been in Secondary and Postsecondary Schools?" Schoen's review contained within it a subset of studies in which the researcher had investigated individualized (self-paced) mathematics programs at the secondary level (pp. 353-354).

This was an appropriate set of studies to use to investigate (by applying the aggregate chi-square methodology)

the true strength of individualization as an educational strategy. First, it was appropriate because Schoen (1976) specifically identified self-pacing (p. 352) as a common factor in all of the studies which he had included in the review. Self-pacing was the key element because self-pacing has been identified as "the only common universal attached to the term . . ." individualization (Good et al., 1975, p. 169).

The fact that self-pacing was a unifying factor supported the requirement that there be a rational basis for combining the studies. The fact that all of the studies included in Schoen's review of research were further narrowed to include only secondary level studies and only mathematics programs helped to support the rational basis for their inclusion in an aggregate chi-square methodology.

Second, Schoen's review was appropriate because each study included in the review had many characteristics which have been commonly identified with the individualized model and thus support (in addition to self-pacing) the rational basis for establishing them as a set to be pooled by the chi-square model.

> First, they were based on a specific set of behavioral objectives. Second, the mathematics content to be learned was divided into small modules or units. Third, learning packets were written for each unit; the learning packets served as guides for the students, enabling them to proceed more or less independently through the content at their own pace. Fourth, for the most part the students learned independently from

> textbooks and worksheets, through some programs included other media. Fifth, each packet contained pretests and posttests: the student was required to pass one or both before proceeding to the next unit. The teacher's role was that of manager, record keeper, individual tutor, and sometimes curriculum developer. (Schoen, 1976, p. 352)

The five general characteristics just quoted plus the self-pacing characteristic established the fact that the studies included in Schoen's review contained studies which were in fact models of individualization and thus met the rational basis requirement.

Third, the studies included in Schoen's review were appropriate because of the definite ambiguity of the results. Of the twelve studies included in the review, one (Bull, 1971) concluded that students studying mathematics with the individualized approach scored better than students studying with a traditional approach. Conversely, Fisher (1973), Herceg (1972), and Hirsch (1972) all concluded that the traditional approach was better.

In contrast to the four previously named researchers who found a significant difference in mathematics achievement depending on the method of instruction used, eight researchers reported no significant difference in either method (Englert, 1972; Hanneman, 1971; Ludeman, 1973, Penner, 1972; Schoen & Todd (Note 2); Schoen & Todd, 1974; Taylor, 1971; Thomas, 1972).

Fourth, the general quality of the research designs and statistical analyses were of high quality. Random assignment of students or classes to treatments was the basic way in

which group equivalency was obtained. Analysis of covariance and analysis of change from a pretest to a posttest score were used to support random assignment in attempting to insure pre-treatment group equivalency (Schoen, 1976, p. 353).

The criterion measure used was either a standardized test or a teacher-made test. Many of the studies measured criteria other than but including achievement. Attitude toward the content being taught or learned was the criterion most often measured other than achievement.

Fifth, another potentially important factor was that the collection of studies in Schoen's review was relatively recent. Bull's 1971 study and Schoen's 1974 study represent the extremes on a time continuum.

Sixth, the studies were relatively easy to acquire. Nine of the studies were doctoral dissertations and were readily available through University Microfilms of Ann Arbor, Michigan. One study was reproduced by the United States Educational Resources Information Center (ERIC).

Seventh, by using the self-contained set produced by Schoen (1976), this researcher was able to avoid the researcher bias (Gage, Note 1, p. 13) of choosing a self-selected set. Only one study from the Schoen review was not used and that was due to its unavailability and not to any arbitrary decision by this researcher.

Eighth, the set of studies reported on by Schoen provided enough independent samples so that the results of

this investigation would be defensible. That is, there were enough separate measures included in the pooled data to insure reliability of the aggregate chi-square statistic. Fisher (1948, p. 100) used 3 independent samples in arriving at an aggregate chi-square statistic. Gage (Note 1, p. 33) used 18 independent samples and Gordon et al. (1952, p. 315) used 13 independent samples. In this investigation, 24 independent samples were used to arrive at the aggregate chi-square statistic.

As was noted earlier in this chapter, Schoen and Todd had two separate investigations which were both done in 1974. Only one of Schoen and Todd's (Note 2) studies was available in a form which included group means. The Schoen and Todd (Note 2) study with group means was available from the authors. The second Schoen and Todd study which was included in the "Research Reporting Sections, Annual Meeting of National Council of Teachers of Mathematics (1974)" has not been used in this study. The group means were not available in the ERIC document or from the authors.

Summary

The main criterion in choosing a set of studies for use with the aggregate chi-square model was that of establishing the rationale basis for inclusion of each study. In this case, the rational basis for inclusion was the characteristic of self-pacing which was present in each study. Self-pacing was focused on because it is the only common

element found in individualized programs (Good et al., 1975, p. 169).

In addition, there were five other characteristics of individualization cited by Schoen and quoted above which supported the establishment of a rational basis for inclusion in the set.

Ambiguity of results was another feature associated with the studies reported in Schoen's review. Ambiguity of results was consistent with other research done on individualized (self-paced) instruction. Therefore, Schoen's collection of studies provided an adequate set on which to apply the aggregate chi-square model.

The studies in Schoen's review had been done over a relatively short time span, were narrowly focused in terms of grade level and subject matter (when compared with other reviews), and were relatively easy to acquire in terms of time and cost.

CHAPTER III

REPORTED RESULTS OF THE INDIVIDUAL STUDIES AND CONVERSION TO CHI-SQUARE

Introduction

Chapter III is a presentation of each of the 11 individual research reports with their statistical results and the aggregate statistic derived by pooling the individual statistics.

Six elements of each study were identified as being appropriate to include. The 6 elements are the background data, the characteristics of individualization, a definition of experimental and control groups as it applied to each of the studies, the hypothesis tested, the reported statistical results, and the conversion to chi-square of the reported results.

In some of the studies more than one hypothesis was being tested. Where it was appropriate (and it was appropriate in all but one case), only one hypothesis and the statistical results for the test of that hypothesis were reported. In the one case (Hirsch) where each of two hypotheses was appropriate to be included, the results of each test of the hypothesis was reported as an independent sample.

Procedures for carrying out the conversion to chi-square varied. The variation in procedure depended on two factors. One factor was whether the hypothesis was stated as directional or non-directional. A non-directional hypothesis required that the probability derived from the test of the hypothesis be divided by two.

A second factor that had to be considered was whether the reported statistical result was in the same direction as the majority of the cases in the set of results being pooled to determine the aggregate chi-square.

The set of independent results used was divided in terms of favoring experimental or favoring control as shown in Table 1.

Table 1

Delineation of Reported Results According to Favoring Experimental or Control Groups

Experimenter	Experimental	Control
Bull (1971)	1	
Englert (1972)	1	2
Fisher (1973)		1
Hanneman (1971)	1	4
Herceg (1972)	1	1
Hirsch (1972)		2
Ludeman (1973)	2	
Penner (1972)	3	2
Schoen & Todd (1974)		1
Taylor (1971)		1
Thomas (1971)	1	
Totals	10	14

The fact that the majority of the results (as shown in Table 1) favored the control groups was an important result and was a determining factor in deriving the aggregate chi-square statistic (refer to Procedures section of Chapter 1). The direction of the results reported as "non-significant" was determined by comparing group means reported in each researcher's original data.

The following sections are in alphabetical order by the author's last name.

Bull (1971)

Background Data

Bull conducted a semester-long study in which he compared the achievement of two individualized geometry classes against the achievement of two comparable classes taught in a traditional manner.

Bull did the study at Camelback High School, Phoenix Union High School District, Phoenix, Arizona.

Characteristics of Individualization

The characteristics of individualization were self-pacing, use of behavioral objectives, and self-choice of learning experiences.

Experimental and Control Groups

The individualized groups were classed as experimental. There were 34 students in each of the two experimental groups and in each of the two control groups. Two teachers taught both an experimental and control group. A two-by-two factorial design was used to analyze the effects of

individualization and time of day. Students were randomly assigned to the respective groups.

Hypothesis Tested

The hypothesis tested was stated by Bull in the following manner.

> There is no difference in the mean test score of geometry students taught by the traditional method and the mean test score of geometry students taught by the individualized instruction method as measured by the Mid-Year Geometry Test. (p. 11)

Statistical Results

Bull applied the *t*-test to his non-directional hypothesis and obtained the result $t = 3.229$.

Conversion to Chi-square

By applying the statistical techniques and extended table of chi-square for two degrees of freedom outlined in Gordon et al.. (1952), a chi-square of .0201 was obtained.

The transformation was as follows.

$$t(67) = 3.229,\ p < .05$$
$$p = 1 - (.001/2) = .9995$$
$$\chi^2 = .0201$$

Bull reported the *t* ratio to have been significant in favor of the experimental group.

Englert (1972)

Background Data

Englert's study was conducted at the Cleveland Heights High School in Cleveland Heights, Ohio. The subjects used in the study were first-year algebra students. The first-year algebra students were classified as low-achievers and

the group was comprised of pupils from grades ten through twelve. The duration of the study was one semester.

Characteristics of Individualization

Englert defined individualization in the following manner. "The emphasis in this approach is upon the individual as he learns and proceeds at his own pace" (p. 9).

Experimental and Control Groups

The experimental group was that group of students which proceeded through the text book at its own pace. There were three experimental groups taught by three different teachers. For comparison, three control groups were established. Each control group was taught by one of the experimental group teachers. Thus, three teachers each taught an experimental and a control group for a total of six groups (pp. 9-10).

Hypothesis

Englert hypothesised that "there is no difference in algebra achievement levels between senior high school students taught by the group-oriented approach" (p. 51).

The standardized Seattle Algebra Test was used as the criterion measure. The hypothesis was non-directional and tested at the .05 level.

Statistical Results

A comparison of group means from the arithmetic pretest to the algebra posttest yielded the following results.

Table 2

Group Means and t-scores for Each Independent Sample

Teacher	Experimental	Control	df	t
A	18.05	22.10	39	2.05*
B	22.33	21.30	42	- .59
C	20.67	21.13	39	.23

Note. *$p < .05$.

Conversion to Chi-square

Each of the teachers represented an independent sample as defined in the Procedures section of Chapter I. Therefore, each chi-square generated from each independent sample (teacher A, teacher B, & teacher C) was entered in the aggregate chi-square table as separate entries.

The original probability of each independent sample was divided by two ($p/2$) because the hypothesis was non-directional.

The independent probability associated with teacher B was subtracted from one ($1-p$) because the result was not in the direction of the majority of the cases in the total set used for the aggregate chi-square statistic. The direction was determined by comparing the group means in the original data as shown in Table 2. Table 3 shows the conversion to chi-square of the reported results.

Table 3

Conversion to Chi-square of Each Independent Sample

Teacher	t	Exact One-tailed Probability	χ^2
A	2.05	$p = .05/2 = .025$	7.3778
B	- .59	$p = 1-(.55/2) = .725$	.6432
C	.23	$p = .80/2 = .40$	1.8326
		Total	9.8536

Fisher (1973)

Background Data

Fisher studied the difference in effect on achievement in geometry when using an individualized approach versus a lecture-demonstration approach. Fisher's study was conducted at Albert Einstein Senior High School, Montgomery County, Maryland. The duration of the study was one year. The subjects in the study were eleventh grade geometry students.

Characteristics of Individualization

The computer was used (computer managed instruction) to supply daily monitoring of each student's progress through his program. The individualized design used in Fisher's study permitted "a student to progress through the material at his own rate" (p. 81).

Experimental and Control Groups

The sample was divided into two groups. One group (experimental) consisted of all students studying geometry through the computer managed instruction plan. The control

group consisted of all the students studying geometry by a traditional geometry curriculum. The criterion measure was the standardized Cooperative Mathematics Test for Geometry.

Hypothesis

Fisher tested the following directional hypothesis.

> The computer-managed behavioral objective instructional curriculum is more effective than the traditional curriculum in developing the basic skills, concepts, and logical reasoning skills of geometry as measured by the Cooperative Mathematics Test for Geometry. (pp. 96-97)

Statistical Results

Fisher's reported statistic of $t(81) = -1.28$, $p > .05$ was not significant. However, when the group means were compared, the control group mean was greater (control group mean 12.23 & experimental group mean 10.77) than the experimental group mean. Therefore, the results were taken as favoring the control group.

Conversion to Chi-square

The calculated value $t(81) = -1.28$, $p > .05$ was taken directly to the extended table of chi-square for two degrees of freedom (Gordon, et al., 1952). An exact one-tailed probability of .20 was obtained. The probability of .20 yielded a chi-square of 3.2189.

The extended table was used directly in this case because the test was one-tailed and the results favored the control which was true for the majority of the independent samples used to derive the aggregate chi-square statistic.

Hanneman (1971)

Background Data

Hanneman conducted his study at Mankato (Minnesota) High School. The duration of the experimental study was 14 weeks. The subjects of the study were tenth-grade geometry students.

Characteristics of Individualization

Individualization was characterized by self-pacing, self-testing (pre and post) and learning activity packages (LAP's). End-of-unit teacher designed tests were used as the criterion measure (dependent variable).

Experimental and Control Groups

The experimental groups were made up of 45 tenth-grade geometry students. The control groups were made up of 47 tenth-grade geometry students. The five experimental groups were exposed to the program of individualization as described in Characteristics of Individualization. The five control groups remained in traditional classrooms. Traditional classrooms were characterized by lecture-discussion and whole group pacing (all students received, performed, and were tested on assignments at the same basic intervals).

Hypothesis

Hanneman's hypothesis was stated in the following manner. "The performance on end-of-unit tests of students receiving instruction through independent study will not differ from the performance of those receiving group instruction" (p. 25).

A two-tailed t-test was used as the measure of the non-directional hypothesis.

Statistical Results

Hanneman organized the experimental and control groups around five learning activity packages (LAP's). The LAP's were designed to be equivalent. The statistical results were reported as independent samples in terms of each of the five LAP's. Table 4 shows the reported results.

Table 4

Group Means and t-scores for Each Independent Sample

LAP	Control	Experimental	t
1	86.3	84.3	.870
2	80.0	76.5	1.357
3	79.9	71.8	2.769*
4	79.4	76.5	1.184
5	84.7	85.1	- .194

Note. * <.01, df was 90 for all LAP's

Conversion to Chi-square

Table 5

Conversion to Chi-square of Each Independent Sample

LAP	t	Exact Probabilities	X^2
1	.870	.40/2 = .200	3.2189
2	1.357	.20/2 = .100	4.6052
3	2.769	.01/2 = .005	10.5966
4	1.184	.30/2 = .150	3.7942
5	- .194	.87/2 = .435 (1-.435)=.565	1.1419
		Total	23.3568

Note. The probability for LAP 5 was subtracted from 1 because the experimental group scored higher than the control group (refer to Table 4).

Herceg (1972)

Background Data

Herceg investigated individualization by comparing the achievement of 16 Algebra 2 classes. Three top-track and 13 middle-track Algebra 2 classes were randomly assigned to three treatment groups. The investigation was conducted in the Gateway School District at Monroeville, Pennsylvania. The duration of the study was not precisely stated.

Characteristics of Individualization

The students in Group A studied Algebra 2 by using computers, behavioral objectives, and were in an "individual rate of learning setting" (p. 49).

Experimental and Control Groups

Group A students used the computer in an individual rate of learning setting with formally presented behavioral objectives. Group A was designated as the [experimental] group. The [control] group, Group B, also used computers and formally presented behavioral objectives, but Group B students remained in a traditional classroom setting.

Hypothesis

Herceg stated the hypothesis in the following manner.

> Students in an individual rate of learning setting who are aware of the bahavioral objectives for a unit in CAM [Computer Assisted Mathematics] will score as high as or higher than students in a traditional classroom setting who are aware of the behavioral objectives for the same Computer Assisted Mathematics unit. (p. 49)

Top track students and middle track students were tested on the hypothesis (p. 49 & p. 54). Therefore, the top-track and middle-track groups of students represent independent samples. The hypothesis was stated as a directional hypothesis; therefore the exact probabilities will not be divided by two.

Statistical Results

Table 6

Group Means and t-ratios For Each Independent Sample

Group	Individualized	Control	df	t
Middle-Track	54.00	61.96	94	5.92*
Top-Track	62.29	62.15	35	0.06**

Note. The hypothesis was tested at $p < .05$.
*$p < .001$
** Not significant

Conversion to Chi-square

Table 7

Conversion to Chi-square of Each Independent Sample

Group	t	Exact Probability	χ^2
Middle-Track	5.92	.001	13.8155
Top-Track	0.06	.90(1-.90)=.10	
		Total	18.4207

Note. $p = .90$ was subtracted from 1 because the experimental group mean was greater than the control group mean (refer to Table 6).

Hirsch (1972)

Background Data

Hirsch compared the effects of guided discovery and individualized instruction on several outcome measures. The duration of the study was one semester. The subjects were tenth-grade Algebra 2 students. The schools were located in Cedar Rapids and Iowa City, Iowa.

Characteristics of Individualization

Individualization was characterized by the use of learning activity packages and progression through the material at each student's own rate (p. 60).

Experimental and Control Groups

Three intact groups of students were assigned three treatments. The three treatments were specified as guided discovery, instructional packages (expository format), and instructional packages (programmed format).

The treatments specified as instructional packages were each individualized as defined above. The instructional package treatments differed "only with respect to programming style" (p. 64). Self-pacing was a key characteristic of both instructional package treatments. Self-pacing was not a characteristic of the guided discovery treatment.

Hypothesis

Hirsch developed four hypotheses each one of which corresponded to a single outcome measure. An assumption was made that the outcome measures of initial learning and

retention were sub-measures of academic achievement. According to Hirsch, initial learning "was specifically designed for this study . . . to provide a measure of student achievement" (p. 67).

Hirsch designed the "Retention Test . . . to provide a measure of student retention of complex number concepts" (pp. 71-72). Many educators would agree that "initial learning" and "retention" are measures of academic achievement.

Conversely, for the purpose of this study, it was not assumed that "lateral transfer" and "vertical transfer" were measures of academic achievement.

Therefore, only hypotheses "I" and "IV" are referred to in the section Statistical Results and which are quoted in the following text.

> Hypothesis I: There are no significant differences among the adjusted group initial learning means for the three treatments.
>
> Hypothesis IV: There are no significant differences among the adjusted group retention means for the three treatments. (p. 83)

Statistical Results

Hirsch used the *F*-test to compare the three treatment group's by the group's respective means. Table 8 shows the comparative mean scores and Table 9 shows the obtained values of *F* for the two outcome measures used in the aggregate chi-square statistic.

Table 8

Adjusted Mean Scores
for Three Treatment Groups and Two Outcome Measures

Treatment	Initial Learning	Retention
Guided Discovery	17.87	9.06
Instructional Packages (expository)	15.75	7.92
Instructional Packages (programmed)	15.78	8.09

Note. Guided Discovery was the control (traditional) group and Instructional Package groups were the experimental (individualized).

Table 9

F-ratios for Each Independent Sample

Outcome Measure	df	*F*	*p*
Initial Learning	2/208	9.19	<.01
Retention	2/208	2.76	<.05

Note. Both *F*-ratios favored the control (refer to Table 8).

Conversion to Chi-square

The outcome measures initial learning and retention were considered to be independent samples. The conversion to chi-square of the *F*-ratio for each outcome measure is as follows.

Table 10

Conversion to Chi-square of Each Independent Sample

Outcome Measure	F	Exact Probability	χ^2
Initial Learning	9.19	.0005/2 = .00025	12.0238
Retention	2.76	.0250/2 = .0125	8.7657
		Total	20.7895

Note. The exact probabilities were divided by 2 because the hypothesis was non-directional.

Hirsch tested for post-hoc comparisons among means and found a significant difference between the control and each experimental adjusted group mean. No significant difference was found between the two experimental group means and thus the reported F-ratio reflects the difference between the control and experimental groups.

Ludeman (1973)

Background Data

Ludeman investigated the effects of an individualized program on ninth-grade algebra and basic mathematics students. The investigation was carried out at the Arnold, Nebraska, public schools. The investigation took place over the course of one school year.

Characteristics of Individualization

Individualization was based on a video-tape program and continuous progress format. Individualized student-teacher-parent contracts were developed and were the basis of a self-paced schedule (p. 16).

Experimental and Control Groups

In both the Basic Mathematics and Algebra I classes, the experimental groups were those that received individualized mathematics and algebra instruction supplemented by video-taped presentations.

The control group was the previous year's ninth-grade class. Ludeman explained that this was due to the fact that Arnold Public Schools were a small rural district. Consequently, there were not enough students to form an experimental and control group from among the current enrollment of ninth-graders. The previous year's ninth-grade class had not received the individualized instructional program.

Hypothesis

There were no written hypotheses for Ludeman's project. An objective was established which was to increase achievement levels by 5-10 percentile points (p. 1), but an hypothesis was not generated or if it was it was not in the reported material.

Statistical Results

For Basic Mathematics, Ludeman reported a control group mean of -1.667 and an experimental group mean of 2.333.

The results yielded $t(22) = 1.580$, $p > .05$. While the t-ratio was not significant the mean scores did favor the experimental.

For Algebra I (after 20 weeks of instruction), Ludeman reported a control group mean of 26.90 and an experimental

group mean of 30.50. The statistic generated from those means was $t(29) = 1.858$, $p > .05$. Again, the t-ratio was not significant, but the mean scores favored the experimental group.

Conversion to Chi-square

Table 11

Conversion to Chi-square of Each Independent Sample

Group	t	1-p	χ^2
Basic Mathematics	1.580	1-.20 = .80	.4463
Algebra 1	1.858	1-.10 = .90	.2107
		Total	.6570

Note. The exact one-tailed probabilities were subtracted from one (1-p) as per Procedures section of Chapter I.

An assumption was made that Ludeman, in calculating the t-ratio, used the one-tailed distribution of t. As was stated earlier, Ludeman reported no written hypothesis. However, when the original data which he did report ($t(22) = 1.580, p > .05$ & $t(29) = 1.858, p > .05$) were compared with the critical values which he also reported (2.074 & 2.045, respectively), the conclusion was reached that Ludeman had worked from the assumption of a one-tailed hypothesis.

Penner (1972)

Background Data

Penner studied the effect of individualization on achievement in trigonometry. The subjects used in the study

were seniors in the public schools of Omaha, Nebraska. The study was conducted during the first semester.

Characteristics of Individualization

Penner used the term "individual progress approach" (p. 4) and stated that this meant an approach "which allowed the students, with the aid of a syllabus, to progress at their own rates" (p. 4).

Experimental and Control Groups

There were five experimental classes (individualized) and five control classes. Five schools participated in the study with a total of 212 students. Each participating teacher taught one experimental and one control class. The students in the experimental groups progressed at their own rates while the students in the control groups were taught in a traditional manner.

Hypothesis

Penner stated the hypothesis in the following manner.

> There is no significant difference in student achievement in trigonomentry between those students who use the individual progress approach and those who use the traditional approach. (p. 5)

Statistical Results

Penner used each of five schools (Benson, Burke, Central, North, and South) as independent samples. Table 12 shows the reported results.

Table 12

Group Means and t-scores for Each Independent Sample

School	Experimental	Control	t
Benson	18.826	17.903	-.545
Burke	20.364	20.192	-.098
Central	25.211	21.588	-2.073
North	13.538	16.080	1.409
South	15.235	25.158	4.331

Conversion to Chi-square

Table 13

Conversion to Chi-square of Each Independent Sample

School	t	Exact One-tailed Probability	X^2
Benson	-.545	1 - (.60/2) = .7000	.7133
Burke	-.098	1 - (.90/2) = .5500	1.1957
Central	-2.073	1 - (.05/2) = .9750	.0506
North	1.409	.175/2 = .0875	4.8723
South	4.331	.001/2 = .0005	6.9078
		Total	13.7397

Note. The exact probabilities were subtracted from 1 because the group means for those schools (Benson, Burke, & Central) favored the experimental.

Schoen and Todd (1974)

Background Data

Schoen and Todd investigated two questions.

> (a) Does the detailed preparation of a concept centered individualized learning package (ILP) by a teacher improve the teacher's ability to teach that concept using either a teacher centered (TC) approach or a learning package approach (as measured by student achievement)?
>
> (b) Is there a difference in student achievement scores on concept taught by ILP as compared to a TC approach? (p. 2)

The results of the investigation of the second question "b" are the results reported on in this study.

Characteristics of Individualization

The characteristics of individualization included the use of behavioral objectives, self-testing on pretests and posttests, learning activity packages, and individual progress through the learning activity packages (p. 3).

Experimental and Control Groups

Six mathematics teachers were paired on the basis of comparability of classes.

> Each teacher in each pair had two classes of mathematics at a certain level and his 'mate' had two classes at the same level. In particualr, each teacher in pair one taught two ninth grade General Mathematics classes, pair two taught two Algebra 1 classes and the third pair taught eighth grade General Math classes. Thus, six teachers and twelve classes were involved in the experiment. (p. 2)

The treatment for the experimental group was the use of ILP's and the control group teachers used a "lecture-discussion" (p. 3) methodology.

Hypothesis

Schoen and Todd did not write out their hypothesis in a standard way. However, from the following quote, the hypothesis can be discerned. "The hypotheses of no differences in achievement scores on the main effects-TC vs. ILP and preparer vs. non-preparer could not be rejected" (p. 12). The hypothesis was clearly null and non-directional in nature.

Statistical Results

The statistic reported by Schoen and Todd was F (1/20) = .08 >.05. Schoen and Todd did not treat each teacher as an independent sample. Therefore, the statistical results were reported by Schoen and Todd as a composite F-ratio. The F-ratio was taken as favoring the control groups because, when the group means were compared for each teacher (Table 14), the control groups had higher scores.

Table 14

Group Means for Schoen and Todd

Teacher	Experimental	Control
1		
Unit 1	7.7	9.5
Unit 2	8.4	8.7
2		
Unit 1	7.4	6.9
Unit 2	7.5	8.1
3		
Unit 1	14.3	12.4
Unit 2	13.2	11.2
4		
Unit 1	13.3	15.7
Unit 2	15.2	15.2
5		
Unit 1	19.0	21.0
Unit 2	34.5	34.5
6		
Unit 1	15.0	11.1
Unit 2	28.7	31.1

Note. Six of the group means favored the control and four favored the experimental while two were identical.

Conversion to Chi-square

$$F(1/20) = .08$$
$$p = .99/2 = .495$$
$$\chi^2 = 1.4064$$

The probability for $F(1/20) = .08$ was not recorded in a standard table. Therefore, the conservative probability of .99 was used.

Taylor (1971)

Background Data

Taylor conducted the study at Crestmoor High School in San Bruno, California. The duration of the study was one semester. The subjects were ninth, tenth, and eleventh graders who were studying Algebra 1.

Characteristics of Individualization

Independent study was defined to be a learning situation in which the students studied alone or in small groups with a minimum amount of help from the teacher. The students used a conventional textbook and progressed individually by completing assignments and tests associated with a given chapter before continuing to the succeeding chapter. "The students progress at their own rate through a specified course of study" (p. 10).

Experimental and Control Groups

There were 48 students used as subjects in the study. The students were enrolled in two classes of Algebra 1. Twenty-five students were in the individualized class and 23 were in the lecture-discussion class. The experimenter taught both classes. Assignments to classes were made by a flip of a coin.

Hypothesis

Taylor generated a non-directional null hypothesis which was stated as follows. "There is no difference

between lecture-discussion and independent study with respect to growth in achievement in Algebra 1" (p. 13).

Statistical Results

The reported results showed that $t(22) = 1.1648$, $p > .05$. The result favored the control group but was not significant. The control group mean was 9.600 while the experimental group mean was 7.522.

Conversion to Chi-square

The reported results showed that $t(22) = 1.1648$. In converting the t-ratio to an exact one-tailed probability, the following result was obtained.

$$p = .25/2 = .125$$

A probability of .125, when taken to the extended table of chi-square, yielded a chi-square of 4.1589.

Thomas (1971)

Background Data

Thomas studied the effects of individualization on five classes of Advanced Algebra students in two high schools of the Lincoln Public Schools at Lincoln, Nebraska. The research covered a period of one school year.

Characteristics of Individualization

The individualized algebra program was characterized by the use of learning activity packages (LAP's) and self-pacing. Thomas specifically describes the self-pacing feature as "an opportunity for each student to proceed at his own rate of speed, commensurate with ability, interest, and motivation" (p. 5).

Experimental and Control Groups

The experimental group was divided into five classes with a total of 102 students. Two teachers were assigned to teach the five classes of continuous progress mathematics (experimental group). The five classes of traditional advanced algebra were taught by three different teachers. There were 122 students enrolled in the traditional classes.

Hypothesis

Thomas stated the hypothesis in the null form and as non-directional in the following manner. "There is no significant difference between the achievement posttest mean of the continuous progress advanced algebra classes and the achievement posttest mean of the traditional advanced algebra classes" (p. 6).

Statistical Results

The results were reported as not being significant. However, the reported group means indicated that the experimental group achieved more than the control group. The experimental group mean was 29.91 and the control group mean was 28.35.

The derived statistic was $F(1/149) = .203$, $p > .05$.

Conversion to Chi-square

$$F(1/149) = .203$$
$$p = .99$$
$$\text{Reversed } p = 1-(.99/2) = .505$$
$$\chi^2 = 1.3664$$

The probability for $F(1/149) = .203$ was not recorded in a standard table. Therefore, the conservative probability of .99 was used.

Computation of the Aggregate Chi-square

Table 15 shows the individual chi-square statistics and the aggregate chi-square statistic. By pooling the individual research results, this researcher was able to use the data from the studies previously reported as not significant. This feature of the aggregate chi-square was very important in that 8 of the 11 original studies were reported as not significant. By pooling the previously reported non-significant studies into the aggregate chi-square, the non-significant studies added data that produced a highly significant result.

Table 15

The Aggregate Chi-square Statistic

Experimenter	Nr. of Independent Samples	Total of χ^2
Bull (1971)	1	5.0201
Englert (1972)	3	9.8536
Fisher (1973)	1	3.2189
Hanneman (1971)	5	23.3568
Herceg (1972)	2	18.4207
Hirsch (1972)	2	20.7895
Ludeman (1973)	2	.6570
Penner (1972)	5	13.7397
Schoen & Todd (1974)	1	1.4064
Taylor (1971)	1	4.1589
Thomas (1971)	1	1.3664
	Aggregate Chi-square	101.9880

Note. Aggregate $\chi^2(48) = 101.9880$, $p < .001$. The df (48) results from multiplying 2 times n (the number of independent samples), (Gordon, Loveland, & Cureton, 1952, p. 314).

The probability .001 had to be adjusted to be consistent with the underlying assumption that the results could

be in either direction (which requires a two-tailed test). Each independent statistical result used to achieve a combined statement of significance was generated from one-tailed probabilities. Even when the original researcher used a two-tailed test, the result had to be converted to a one-tailed probability (as per procedures).

The problem (and therefore the reason why the probability of the aggregate chi-square statistic has to be adjusted) came because of the fact that the aggregate chi-square procedure allowed for a result in either direction. When the procedure allows for a result in either direction, a two-tailed probability is required.

"The tabled significance value must therefore be doubled" (Gordon et al., 1952, p. 315) in order to avoid the contradiction of the one-tailed result (using one-tailed probabilities in the aggregate chi-square) with the two-tailed assumption (willingness to consider a result in either direction). The probability, then, is .002 and not .001.

Summary

The aggregate chi-square statistic for the 24 independent research samples of self-paced individualized instruction showed that self-paced individualized instructional practices were not superior to traditional instructional practices. In fact, the reverse could be implied;

that is, traditional instructional practices were superior to self-paced individualized instructional practices. The probability of .002 for this combined significance test indicated that in 998 of every 1,000 cases this result would be replicated. In other words, it was very unlikely that this result came by chance.

CHAPTER IV

SUMMARY, CONCLUSIONS, IMPLICATIONS, AND SUGGESTIONS FOR FURTHER RESEARCH

Summary

The problem of this study was to determine the effectiveness of self-paced individualized instruction by applying the aggregate chi-square statistical procedure to the results of several, previously done, independent studies which have analyzed self-paced individualized instruction in its relationship to student achievement. What would be the result when these several independent results were pooled to derive a single statistic?

The accountability movement and a philosophical/emotional commitment of professional educators have exerted considerable force on local school administrators to adopt individualization as the dominant (if not only) instructional strategy. This pressure has been very forceful in spite of the fact that empirical research has not supported the claims made for individualized instruction. Empirical research studies done on individualized instruction have presented an inconsistent and therefore an indicisive pattern of results. The aggregate chi-square procedure was applied to the results of a set of 11 previously reported studies in an attempt to discern a consistent pattern.

The strength of the aggregate chi-square procedure lies in the use that was made of non-significant data. Schoen (1976) reported on the results of 12 independent researchers. Of the 11 studies included from Schoen's review in this study, 8 of the researchers reported non-significant results. Unless some secondary analysis was done, the review was not very enlightening or helpful in educational decision making that was concerned with self-paced individualized instruction.

However, when the results of the 8 non-significant studies were able to be included in a statistical procedure such as the aggregate chi-square, a completely new and much stronger picture was obtained. This result was consistent with the kinds of results that Gage (Note 1) obtained when applying the aggregate chi-square procedure to sets of independent studies which had been previously done on several teacher variables. Instead of a capricious and inconsistent picture, a clear indication of the strength of the variables on the outcomes measured was clearly evident. Specifically, in this study, the result was aggregate $\chi^2(48) = 101.9880$, $p < .002$. This result indicated that self-paced individualized instructional practices were not superior to traditional instructional practices. In fact, the reverse could be implied; that is, traditional instructional practices were superior to self-paced individualized instructional practices. The probability of .002 for this combined significance test indicated that in 998 of every 1,000 cases

this result would be replicated. In other words, it was very unlikely that this result came about by chance.

Conclusions

The results of this study indicated that traditional classroom instructional practices were superior to individualized (self-paced) instructional practices. In evaluating this conclusion, it should be recognized that the procedure used in this study did not take into account the relative merit of the individual research studies in terms of design, interaction effects within studies and a host of other technical considerations. The basic assumption was that the independent statistical results generated by each researcher were adequate statistical statements of the strength of self-paced individualized instruction as measured by that study.

Implications

There were two important implications derived from this study. One implication was in the area of research techniques. The emphasis that Gage (Note 1), Glass (1976) and Light and Smith (1971) have placed on secondary analysis of educational research seemed completely justified. The weak and vascilating research results often reported in educational research reviews have not been a true account of the status of educational practices. This study supported the notion that what is done in classrooms is important

and does effect a student outcome such as achievement. The viability of the aggregate chi-square as a tool for secondary analysis was also supported.

A second implication was in the area of decision making regarding educational programs and practices. Educational leaders have been susceptible to the political, philosophical, and social forces which constantly impinge on the life of the schools. Accountability has been a force that has vitally affected the programs in the schools (Darland, 1970; Davies, 1970; Lessinger, 1971; Morris, 1971; Sciara & Jantz, 1972).

Educational leaders have been pressed to defend their programs and practices with measureable results (Dunkin & Biddle, 1974; Lessinger, 1971). More specifically, they have been pressed to defend their programs and practices with measureable results focused on the achievement of individual students (Davies, 1970; Morris, 1971). The practice of self-paced individualized instruction was a logical response to the demand for measureable results. The progress of individual students was thought to be a more precise gauge of success in the classroom than group norms produced by standardized norm referenced tests.

The program of self-paced individualized instruction was also a logical response to another force that was beginning to make a strong impact on the educational community. As Davies (1970) and Morris (1971) indicated, there was a need to find an educational vehicle through which

and by which the maximum potential of the individual student could be realized. Self-paced individualized instruction seemed to offer the solution to the problem of accountability and to the need for individualized plans of instruction.

It appeared that there had been a happy marriage of the "commitments" to accountability which would be measured by individual student performance and the commitment to the development of "individual human beings" (Davies, 1970, p. 129). Self-paced individualized instruction provided a common ground for these two commitments and practicing educators have felt the persistent weight of the force generated by these two commitments.

The results of this study indicated that self-paced individualized instruction was not the teaching/learning strategy that would satisfy the demands of the accountability movement nor was it able to meet the idealistic demands of developing every individual to his fullest potential (at least not academic potential). Perhaps some other form of individualized instruction could meet those demands and herein lies a hint at one direction further research on the topic of individualized instruction might take.

Suggestions for Further Research

In this study, self-paced individualized instruction was investigated. It might be very helpful if some

empirical research studies were done in which self-paced and teacher-paced individualized instruction were compared. The following comments will serve to strenghten this point and also serve to clarify what is meant by "teacher-paced."

Taveggia (1976) reviewed 14 separate studies each of which compared the "learning outcomes of a new instructional procedure, the 'Personalized System of Instruction' (PSI), with the learning outcomes of conventional approaches to college teaching" (p. 1028). Taveggia made a potentially significant observation when he pointed out that "five features probably account for the superiority of PSI over conventional methods" (p. 1030). The Personalized System of Instruction was superior to conventional methods of college instruction in all 14 studies reported in Taveggia's review.

The significant point made by Taveggia was that the feature termed "go-at-your-own-pace" (p. 1030) was one of five features that seemed to make the PSI appraoch better than conventional approaches. Also, that the term go-at-your-own-pace was a misnomer. "A more appropriate designation would be monitored pacing or forced pacing" (p. 1030). The following excerpt presents a more complete picture of the forced pacing concept of the PSI approach.

> A second, less obvious option suggested by the explanation developed above for the superiority of PSI is to reorganize one's conventional courses, "building in" the unit-perfection, forced-pacing, and monitored progression features of PSI. The

> available evidence suggests that these probably are the features which account for the superiority of PSI over conventional methods. Thus, to the degree that these features are incorporated into conventional courses, student mastery of course content material probably will be enhanced. (Taveggia, 1976, p. 1031)

The concepts of unit-perfection, forced-pacing, and monitored progression mentioned above, do indeed point to a definite forcing of the pace which is in direct contrast to the concept of self-pacing.

There does seem to be a significant relationship between what kind of "pacing" is used and the strength of an individualized methodology in terms of that methodology's ability to change student outcomes. While it was recognized that the studies reviewed by Taveggia were at the college level, the results may have important implications for individualized instructional techniques at all levels and may be revealing as to why self-paced individualized instruction did not fare well when submitted to the aggregate chi-square procedure. In the future, when studies of individualized instruction are conducted, there should be an effort to more clearly define the "pacing" function and to attempt to establish its direct relationship to student outcomes.

APPENDIX

HISTORY OF AGGREGATE CHI-SQUARE AND TRANSFORMATION PROCEDURES

The aggregate chi-square procedure was developed by Fisher (1948). The aggregate chi-square procedure was based upon the fact that the distribution of the sum of several values of chi-square was itself distributed as a chi-square for two degrees of freedom was -2 times the natural logarithm of the probability (pp. 99-101).

While other researchers have discussed the technique and its possible applications (Lancaster, 1949 & Wallis, 1942), there have been few attempts to use the aggregate chi-square method in published research studies. Perhaps the reason for the few applications of the aggregate chi-square has been due to the fact that an extended table of values for chi-square with two degrees of freedom was not available until 1952. It was in 1952 that Gordon et al. developed the extended table which could be used in combining probabilities from independent samples. However, it has to be noted that there were no published studies known to this researcher using the aggregate chi-square even after the Gordon et al. (1952) extended table was published. There was, of course, one exception to the previous assertion and that exception was Gage (1977).

The chi-square model was based on the proof that any p value could be transformed to a chi-square value with two degrees of freedom and that the sum of independent chi-squares was distributed as a chi-square. Following is an outline of the transformation procedure as developed by Fisher (1948, pp. 99-101).

1. The transformation equation:

$$\chi^2 = -2 \log_e p \qquad (1)$$

2. Composite χ^2 is given by the formula:

$$\chi^2 = -2 \sum_{i=1}^{k} \log_e p_i \qquad (2)$$

3. Degrees of freedom:

   $2k$ degrees of freedom where k is the number of independent probability values to be combined.

4. Joint probability of k independent results:

   The product of the k separate p values.

REFERENCE NOTES

1. Gage, N. L. Four cheers for research on teaching. Unpublished research report, 1976. (Avaliable from N. L. Gage, College of Education, Stanford University, Palo Alto, California, 94305). The Gage report was later published and was added to the reference section of this study.

2. Schoen, H. L., & Todd, R. M. Teacher prepared learning packages: Aid to student? or teacher? Unpublished research report, 1974. (Available from Harold L. Schoen, W104A East Hall, University of Iowa, Iowa City, Iowa, 52242).

REFERENCES

American Educational Research Association. Second report of the committee on the criteria of teacher effectiveness. Journal Educational Research. 1953, 46, 641-658.

Anderson, R. H. Individualization the unfulfilled goal. Educational Leadership, 1977, 34, 323-234.

Berry, M. Educational research and the tension between politics and science. Educational Researcher, 1977, 6, (5), 3-6.

Block, J. H. Individualizing instruction: A mastery learning perspective. Educational Leadership, 1977, 34, 337-341.

Briner, C. Administrators and accountability. Theory Into Practice, 1969, 8, 203-206.

Bull, S. S. A comparison of the acheivement of geometry students taught by individualized instruction and traditional instruction (Doctoral dissertation, Arizona State University, 1971). Dissertation Abstracts International, 1971, 31/09-A, 4616. (University Microfilms No. 71-5976)

Darland, D. D. The profession's quest for responsibility and accountability. Phi Delta Kappan, 1970, 52, 41-44.

Davies, D. The relevance of accountability. Journal of Teacher Education, 1970, 21, 127-133.

Dunkin, M. J., & Biddle, B. J. The study of teaching. Atlanta: Holt, Rinehart, & Winston, Inc., 1974.

Englert, T. J. A comparative study of the effects on achievement and changes in attitude of senior high school students enrolled in first-year algebra under two different teaching approaches (Doctoral dissertation, Cornell University, 1972). Dissertation Abstracts International, 1972, 33/03-A, 1076. (University Microfilms No. 72-23, 657)

Fisher, M. E. A comparative study of achievement in the concepts and fundamentals of geometry taught by computer-managed individualized bahavioral objective instructional units versus lecture-deminstration methods of instruction (Doctoral dissertation, The George Washington University, 1973). Dissertation Abstracts International, 1973, 34/05-A, 2161. (University Microfilms No. 73-25, 330)

Fisher, R. A. Statistical methods for research workers (2nd ed.). New York: Hafner Publishing Company, 1948.

Fox, D. J. The research process in education. New York, Holt, Rinehart, and Winston, 1969.

Gage, N. L. Scientific basis of the art of teaching. New York, Teachers College Press, 1977.

George, T. W., & Maxwell, A. The validation of individualized instruction: A glance at an iceberg. Educational Technology, 1976, 16(2), 56-57.

Glass, G. V. Primary, secondary, and meta-analysis of research. Educational Researcher. 1976, 5(10), 3-8.

Good, T. L., Biddle, B. J., & Brophy, J. E. Teachers make a difference. Atlanta: Holt, Rinehart, & Winston, 1975.

Gordon, M. H., Loveland, E. H., & Cureton, E. E. An extended table of chi-square for two degrees of freedom, for use in combining probabilities from independent samples. Psychometrika, 1952, 17, 311-315.

Hanneman, J. H. An experimental comparison of independent study and conventional group instruction in tenth-grade geometry (Doctoral dissertation, University of Florida, 1971). Dissertation Abstracts International, 1972, 32, 11A. (University Microfilms No. 73-4133).

Herceg, J. A study of the coordinator's role in the introduction of formally presented objectives and individualized learning rates in computer-assisted mathematics (Doctoral dissertation, University of Pittsburgh, 1972). Dissertation Abstracts International, 1973, 33/08-A, 4059. (University Microfilms No. 73-4133)

Hirsch, C. R. An experimental study comparing the effects of guided discovery and individualized instruction on initial learning, transfer, and retention of mathematical concepts generalizations (Doctoral dissertation, The University of Iowa, 1972). Dissertation Abstracts International, 1973, 33/07-B, 3194. (University Microfilms No. 73-640)

Hirsch, C. R. A review of research on individual instruction in secondary mathematics. School Science and Mathematics, 1976, 6, 499-509.

Hostrop, R. W., Mecklenbarger, J. A., & Wilson, J. A. (Eds.), Accountability for educational results. Hamden: The Shoe String Press, 1973.

Huck, S. W., Cormier, W. H., & Bounds, W. G. Reading statistics and research. New York: Harper & Row, 1974.

Lancaster, H. O. The combination of probabilities arising from data in discrete distributions. Biometrika, 1949, 36, 370-382.

Lessinger, L. M. Robbing Peter to pay Paul. Educational Technology. 1971, 11(1), 11-14.

Light, R. J., & Smith, P. V. Accumulating evidence: Procedures for resolving contradictions among different research studies. Harvard Educational Review, 1971, 41, 429-471.

Ludeman, C. Final evaluation report, project video-tape packages mathematics. Bureau of Elementary and Secondary Education (DHEW/OE), Washington, C. D., 1973. (ERIC) Document Reproduction Service No. #D 086 545)

Miller, R. L. Individualized instruction in mathematics: A review of research. Mathematics Teacher, 1976, 69, 345-327.

Morris, J. E. Accountability: Watchword for the 70's. The Clearing House, 1971, 45, 323-327.

Penner, H. D. An analysis of using an individual progress approach to the teaching of trigonometry in the Omaha, Nebraska, public high schools (Doctoral dissertation, University of Nebraska, 1972). Dissertation Abstracts International, 1972, 33/94-A, 1421. (University Microfilms No. 72-27, 419)

Saylor, G. Three essentials in determining the quality of education. Educational Leadership, 1977, 34, 243-245.

Schoen, H. L. Self-paced mathematics instruction: How effective has it been in secondary and postsecondary schools? Mathematics Teacher, 1976, 69, 352-357.

Schoen, H. L. Research and its implications for secondary school mathematics instruction via self-pacing. The High School Journal, 1977, 61, 131-140.

Schoen, H. L., & Todd, R. M. Teacher prepared learning packages: Aid to student? or teacher? Research Reporting Sections, Annual Meeting of National Council of Teachers of Mathematics, Atlantic City, New Jersey, 1974. (ERIC Document Reproduction Service No. ED 088 702)

Sciara, F. J., & Jantz, R. K. (Eds.), Accountability in American education. Boston: Allyn & Bacon, 1972.

Taveggia, T. C. Personalized instruction: A summary of comparative research, 1967-1974. American Journal of Physics, 1976, 44, 1028-1033.

Taylor, L. M. Independent study versus presentation by lecture and discussion: A comparative study of attitude and achievement in two algebra 1 classes (Doctoral dissertation, University of Northern Colorado, 1971). Dissertation Abstracts International, 1972, 32/07-A, 3877. (University Microfilms No. 72-3307)

Thomas, B. G. Continuous progress advanced algebra in the Lincoln public schools - A study of achievement and attitude toward mathematics (Doctoral dissertation, University of Nebraska, 1971). Dissertation Abstracts International, 1972, 32/11-A, 6113. (University Microfilms No. 72-16, 019)

Wallis, W. A. Compounding probabilities from independent significance tests. Econometrika, 1942, 10, 229-248.

Webb, L. L., & Howard, T. E. Individual learning: An achievable goal for all! Educational Leadership, 1977, 34, 356-359.

Weber, G. The cult of individualized instruction. Educational Leadership, 1977, 34, 326-329.

BIOGRAPHICAL SKETCH

Paul Ivan Johnson was born October 9, 1937, at Boise, Idaho. In May, 1955, he was graduated from Gooding High School at Gooding, Idaho. In January, 1956, he enlisted in the United States Navy and served four years in the Pacific Area as a Communications Technician. Following his discharge, he entered a training program with a Christian service organization.

In February, 1961, he enrolled at Bethel College in St. Paul, Minnesota, and graduated from that institution with an elementary teaching major in 1967. From 1967 through 1972 he taught at the elementary level in the public schools of Camarillo, California. In 1972 he received the degree of Master of Arts in Educational Administration and Supervision for the elementary school level. The Master of Arts degree was received from San Fernando Valley State College at Northridge, California. From 1972 until the present he has served with the Wycliffe Bible Translators as a school administrator. He currently holds the position of International Coordinator for Children's Education. In 1974 he was granted a study furlough and subsequently enrolled in the Graduate School of the University of Florida, where he began his work toward the degree of Doctor of Philosophy.

Paul Ivan Johnson is married to the former Dona Jean Wilson. He is a member of the American Educational Research Association, Association for Supervision and Curriculum Development, National Association of Elementary School Principals, National Association of Secondary School Principals, and National Educators Fellowship.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ralph B. Kimbrough, Professor
of Educational Administration

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert S. Soar, Professor
of Foundations in Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Michael Y. Nunnery, Professor
of Educational Administration

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Paul S. George, Associate Professor of General Teacher Education

This dissertation was submitted to the Graduate Faculty of the Department of Educational Administration and Supervision in the College of Education and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

June 1979

Dean, Graduate School